Technical Documentation for the 2004 South Carolina Readiness Assessment of Kindergarten and Grade-One Students

South Carolina Department of Education
Inez M. Tenenbaum, State Superintendent

Division of Curriculum Services and Assessment
Lucinda Saylor, Deputy Superintendent

Office of Assessment
Dr. Theresa Siskind, Director
Dr. Jim Casteel, Education Associate
Dr. Imelda C. Go, Education Associate
Joseph C. Saunders, Education Associate

CONTENTS

LIST OF TABLES iii
INTRODUCTION 1

CHAPTER 1. THE SCRA IN 2003–04 3

1.1 SURVEY OF TEACHERS ON THE SCRA 3
1.2 ADVISORY COMMITTEE DELIBERATIONS 4
Composition of the Committee 4
November 2001 Meeting 4
January 2002 Meeting 4
March 2002 Meeting 5
April 2002 Meeting 5
1.3 REVISIONS FROM 2002–03 5
Developmental Guidelines 5
Checklists 6
Reports 7

CHAPTER 2. TEST ADMINISTRATION AND SCORING 8

2.1 OVERVIEW 8
2.2 ADMINISTRATOR TRAINING 8
2.3 SOUTH CAROLINA READINESS ASSESSMENT PROFILE INTERACTIVE 9
History 9
Online System Features 10
Transferring Students 10
2.4 SCORING 11
2.5 SCORE REPORTING 11
Student Home Report 12
Class Roster 12
School, District, and State Summaries 12

CHAPTER 3. TECHNICAL CHARACTERISTICS OF THE INDICATORS AND RATINGS 14

3.1 YEAR-END PERFORMANCE INDICATORS 14
3.2 YEAR-END PERFORMANCE RATINGS 15

CHAPTER 4. VALIDITY 17

4.1 OVERALL VIEW OF DIF ANALYSIS 17
General Description 17
Standardized Mean Difference Procedure 18
NAEP DIF Classification 18
4.2 RESULTS OF DIF ANALYSES 18

APPENDIX A 20
APPENDIX B 23
WORKS CITED 30

TABLES

1.1 2002–03 and 2003–04 SCRA Functional Components for Kindergarten and Grade One....6
2.1 Frequency of Performance Ratings Needed to Achieve Various Component Ratings 13
3.1 Summary of Major Statistics for Year-End Performance Indicators.....15
3.2 Summary of Major Statistics for Year-End Performance Ratings.....15
3.3 State Summary of Year-End Performance Ratings of the Kindergarten Assessment for First-Grade Readiness.....16
3.4 State Summary of Year-End Performance Ratings of the First-Grade Assessment for Second-Grade Readiness 16
4.1 Summary of DIF Classification for Year-End Performance Indicators.....19
4.2 Summary of DIF Classification for Year-End Performance Ratings 19
A.1 Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade.....21
A.2 Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade.....22
B.1 DIF Data: Year-End Performance Indicators for Kindergarten and Gender.....24
B.2 DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity 25
B.3 DIF Data: Year-End Performance Indicators for First Grade and Gender.....26
B.4 DIF Data: Year-End Performance Indicators for First Grade and Ethnicity 27
B.5 DIF Data: Year-End Performance Ratings for Kindergarten and Gender.....28
B.6 DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity.....28
B.7 DIF Data: Year-End Performance Ratings for First Grade and Gender.....29
B.8 DIF Data: Year-End Performance Ratings for First Grade and Ethnicity.....29

INTRODUCTION

The South Carolina Readiness Assessment (SCRA) is a state-level assessment for kindergarten and grade-one students. It replaced the Cognitive Skills Assessment Battery (CSAB), which had been used in the state of South Carolina since 1978. Unlike the CSAB, which was administered in the first month of the first grade and was a single, high-stakes test, the SCRA is a system of continuous assessment that documents a student's performance through focused observations of daily classroom activities. The CSAB was used for the last time in fall 2001.

The development of the SCRA began in 1999 when the South Carolina Department of Education (SDE) awarded a contract to Harcourt Educational Measurement (HEM) to develop a statewide continuous assessment system that would provide a better understanding of student readiness for school. The HEM submitted a proposal to use a version of the Work Sampling System (WSS) modified specifically for the state of South Carolina.

The WSS, used in many states, is a curriculum-embedded continuous assessment process where teachers observe their students during everyday classroom activities to gain a fuller picture of the students' development. Focusing on three key areas—English language arts, mathematics, and personal and social development—these observations can provide a clearer understanding of a child's readiness for school. The modified WSS is aligned with the South Carolina curriculum standards.

The WSS is the major product of Rebus, the company established by Samuel Meisels to develop, market, and support this assessment system. Since 1991, Rebus has been able to accomplish this objective in all fifty states, the District of Columbia, and Canada (HEM 1999, 25). The Department of Defense Education Agency (a civilian agency of the U.S. Department of Defense) adopted the program for all of the kindergarten through third-grade classrooms worldwide. During the 1990s, a number of states, including Pennsylvania and Maryland, started with a small sample of their schools using the system. After initial positive responses in the District of Columbia, for example, administrators in that system decided to implement the WSS in all 116 of their elementary schools over a six-year period (Monrad and Mandeville 1996).

Pearson bought Rebus in 2000. The name Rebus still exists and is mostly associated with the WSS. In 1999, the SDE granted the HEM the contract to develop a readiness assessment for students in kindergarten and grade one. The HEM has contracted with Rebus to use the WSS. However, the HEM is providing the scoring services, using its own in-house facilities.

The SCRA is an ongoing assessment of students in kindergarten and grade one. Its mechanism is a checklist that is completed by teachers on three occasions during the school year. Only the last checklist assessment is used for reporting students' performance. The checklist is organized on the basis of three subject domains, which are labeled "Personal and Social Development," "English Language Arts," and "Mathematics." Each of these domains is divided into a series of functional components under which are sets of performance indicators that are based on the South Carolina curriculum standards. During workshops that are normally conducted in the summer, teachers receive training in completing the checklist according to a set of developmental guidelines.

The first statewide SCRA field test took place during the 2000–01 school year. The system centered on the use of a checklist of developmental indicators that kindergarten and first-grade teachers completed three times during the school year for each of their students. The two checklists, one for students in kindergarten and one for those in grade one, were accompanied by development guidelines designed to give all teachers a similar understanding of the checklist items. Minor revisions were made to the system on the basis of this first pilot test.

The SCRA was field-tested on a larger scale during the 2001–02 year. An SCRA advisory committee was also established in November 2001. This group consisted of teachers and administrators from schools around the state, SDE administrators, and personnel from both the HEM and Rebus, Inc., the company that created the version of the WSS on which the SCRA is based. The advisory committee has continued to meet regularly and has recommended a number of substantial changes to the SCRA. The committee made major revisions for 2002–03, including changes both to the checklists and to the guidelines. The 2002–03 checklists were used in 2003–04. The 2003–04 school year was the last year of the HEM's contract with the SDE.

The first section of this report, chapters 1 and 2, provides an overview of the checklists, the test administration, and the scoring. The second section, chapters 3 and 4, documents the technical characteristics of the test items. Detailed information about the evolution of readiness assessment in South Carolina, the WSS and its being modified to become the SCRA, and technical aspects of previous SCRA administrations can be found in an earlier technical report (Huynh, Prior, and Gallant-Taylor 2004).

Chapter 1

THE SCRA IN 2003–04

This chapter describes how the 2001–02 SCRA teacher survey administered to kindergarten and grade-one teachers and the 2001–02 SCRA advisory committee deliberations affected the 2002–03 SCRA administration. The 2002–03 checklists were used in 2003–04.

1.1 SURVEY OF TEACHERS ON THE SCRA

In an effort to continue their evaluation of the SCRA in prior years, SouthEastern Regional Vision for Education (SERVE) developed and sent surveys to South Carolina kindergarten and grade-one teachers. The survey was developed as a result of teacher focus groups. At the January 2002 SCRA advisory committee meeting, copies of the drafted SCRA teacher survey were made available for committee members' review. Committee members recommended changes in the method of distributing the survey and in the wording of survey items; they also added some survey items. The survey was piloted with teachers in the committee members' respective districts. On the basis of the survey pilot, the committee was to make recommendations to SERVE.

The final version of the SCRA teacher survey consisted of forty items. Thirty-five items were multiple choice, three items were open-ended, and two items were on a continuum scale. The survey was divided into four major sections: "Completing the SCRA" (seventeen items), "Using Data from the SCRA" (six items), "Training and Support for the SCRA" (ten items), and "Respondents to the Survey" (seven items).

In February 2002, the SCRA teacher surveys were sent to districts for distribution to their kindergarten and grade-one teachers. The deadline for the completion and return of surveys was March 22, 2002. There were 4,540 surveys returned. Preliminary results were available in April for the advisory committee's meeting. These results included responses to questions such as these:

- "How often do you think the SCRA rating should be completed during the school year?"
- "What types of documentation do you think teachers should be *required* to keep to support the ratings they give children? Mark all that apply."
- "When do you think it would be most helpful to receive the reports?"
- "The SCRA has impacted my district in the following ways: mark all that apply."

1.2 ADVISORY COMMITTEE DELIBERATIONS

Composition of the Committee

The SCRA advisory committee, which was formed in November 2001, has twenty-five members. Eleven of them are school district personnel: two from the Greenville County School District and one each from Spartanburg School District Three, York School District One, the Berkeley County School District, Florence School District One, the Charleston County School District, Richland School District One, Greenwood School District Fifty, Lexington School District One, and Lexington School District Five. Three are faculty members at the University of South Carolina, Furman University, and Presbyterian College. Four are school teachers: one from Richland School District One, two from Lexington School District One, and one from Florence School District Four. The remaining seven are a representative from the HEM, two representatives from Pearson Early Learning, and four representatives from the SDE. Advisory committee meetings were held November 2001, January 2002, March 2002, and April 2002.

November 2001 Meeting

As stated earlier, the November meeting was the SCRA advisory committee's initial one. Committee members discussed the following topics:

- current student home reports,
- consistency of teacher ratings,
- removing the "progress" rating column from the checklist and the student home report,
- changing the term *proficient* so as not to lead to a comparison to the PACT,
- statewide reporting of SCRA results,
- using electronic checklists, and
- staff development for summer 2002 to train lead teachers for each school.

Three subcommittees were formed to address these three key issues for 2002–03: guidelines revisions, report revisions, and general procedures.

January 2002 Meeting

Each of the subcommittees had at least one facilitator with expertise in that particular area. Two representatives from Pearson Early Learning served in this role for the subcommittee on guidelines revisions. A representative from the HEM assisted with the report revision subcommittee, and a representative from the SDE worked with the general procedures subcommittee.

At the meeting, a representative from SERVE presented a general overview of the procedures and findings of the SERVE review of the SCRA. Draft copies of the teacher survey were provided to committee members. Other topics discussed were definition of ratings, possible approaches to the proposed lead teacher training for summer 2002, and the use of an SCRA alternate assessment.

March 2002 Meeting

The committee reviewed the new guidelines draft from Pearson Early Learning. There was discussion on the inclusion of the state curriculum standards in the guidelines. The committee also discussed the number of student ratings with the SCRA per year and ways to communicate to teachers the changes for 2002–03.

April 2002 Meeting

The advisory committee reviewed the draft copies of the score reports, checklists, and guidelines. The score reports and the checklists contained changes in the wording of the ratings and the rating periods, and the "progress" rating column had been removed. Two versions of the student report (one for teachers and one for parents) had also been developed. Additionally, functional components had been changed to make them consistent with the state curriculum standards in English language arts.

Committee members agreed to have four ratings per year, with only two of them required: "Performance to date (Winter)" and "Year-End Performance." Pearson Early Learning representatives presented notes on a two-day summer workshop agenda for teachers and administrators. Teachers and administrators were scheduled to attend the first day; the second day was designed for teachers only.

1.3 REVISIONS FOR 2002–03

As is stated above, the advisory committee meetings during 2001–02 resulted in revisions to the SCRA. At its initial meeting in November, the committee discussed redesigning the student home report, removing the "progress" part of the checklist, changing the timing of the data collection and reports, and writing new guidelines documents with more specific ties to the state standards.

Developmental Guidelines

The new guidelines reflect the language of the fourth version of the WSS, the 2000 South Carolina mathematics curriculum standards, and the 2002 South Carolina English language arts curriculum standards. The guidelines are issued as one document that covers both kindergarten and grade one. For each indicator, the left-hand page contains the guidelines for kindergarten, and the right-hand page contains the guidelines for grade one. On each page of the guidelines, the SCRA indicator, the related state standards for English language arts and mathematics, the rationale, and examples are given. The guidelines also include an introduction, an explanation of the ratings, and acknowledgements.

Checklists

Changes in the checklists involved the subject domains titles, the functional components, the performance indicators, and the assessment periods.

There were two changes in the titles of the subject domains: the "Language and Literacy" domain was renamed "English Language Arts," and the "Mathematical Thinking" domain was renamed "Mathematics." The title "Personal and Social Development" remained the same.

Table 1.1 reflects the changes in the functional components. The number of functional components was decreased from fifteen in 2001–02 to fourteen for 2002–03. The two functional components "Listening" and "Speaking" were combined under the new functional component labeled "Communication."

Under "Personal and Social Development" domain, the number of indicators for kindergarten and grade one in 2002–03 were decreased to thirteen, compared to sixteen in 2001–02. In the kindergarten "English Language Arts" domain, the number of indicators for 2002–03 was increased to twelve, from eleven in 2001–02. The number of first-grade "English Language Arts" indicators remained at twelve. In the kindergarten "Mathematics" domain, the number of indicators for 2002–03 was increased to fourteen, from thirteen in 2001–02. For the first grade, the number of indicators for "Mathematics" remained at fourteen. After these changes were made, the SCRA development checklists for 2002–03 were comprised of thirty-seven performance indicators for kindergarten and thirty-nine performance indicators for the first grade; and both grade levels had the same functional components for each of the three domains as shown below in table 1.1. The same checklists were used in the 2003–04 school year.

TABLE 1.1

2002–03 and 2003–04 SCRA Functional Components for Kindergarten and Grade One

Domain	Functional Component
Personal and Social Development	A. Self-Concept
	B. Self-Control
	C. Approaches to Learning
	D. Interaction with Others
	E. Social Problem-Solving
English Language Arts	A. Communication
	B. Reading
	C. Writing
Mathematics	A. Mathematical Processes
	B. Number and Operations
	C. Patterns, Relationships, and Functions
	D. Geometry and Spatial Relations
	E. Measurement
	F. Data Collection and Probability

The SCRA mastery levels *not yet, in process,* and *proficient* were changed to the performance indicators *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates*. The descriptions of the assessment periods were changed from "Fall," "Winter," and "Spring," with the overall rating in "Spring," to the following: "Performance to date (Fall)," "Performance to date (Winter)," "Performance to date (Spring)," and "Year-End Performance."

Reports

Changes were also made in the student home report. On the new student home report, all performance indicators will appear. However, there will be two versions of the report, one for parents and one for teachers. The parent version will show the indicators but will have ratings only for the functional components, whereas the teacher version will show the ratings for both the indicators and the functional components. The progress indicators will no longer be on the report.

Chapter 2

TEST ADMINISTRATION AND SCORING

2.1 OVERVIEW

In 2003–04 a teacher made at least two assessments of his or her students during the school year (fall, winter, spring, and end of year) to determine their current level of achievement—*consistently demonstrates*, *sometimes demonstrates*, or *rarely or never demonstrates*—with regard to each of the indicators. The winter and end-of-year assessments were required while the fall and spring ones are optional. Student data were submitted either through machine-scannable checklists or through the South Carolina Readiness Assessment Profile Interactive (SCRAPI) online system. The machine-scannable checklists were sent to the HEM San Antonio scoring center for scoring at the end of the school year. Districts were given the option to submit their data either by using the machine-scannable checklists or by using the online system.

The SCRAPI online system was piloted in 2002–03 (see section 2.3, below), when approximately 25 percent of kindergarten and first-grade teachers used it to submit student data. In 2003–04, approximately 70 percent used SCRAPI. The online system uploads student ratings to a state database so that results are always current and no physical reports need to be submitted.

All students enrolled in kindergarten and grade one for at least one month were assessed. Provisions were made for those students who transferred either into the school or out to another school. For each kindergarten and first-grade student new to the school, a scannable checklist was to be filled out by the receiving teacher if the student arrived within four weeks before the end of the assessment period. The transferring school was required to forward the scannable checklist to the new school as part of the student's permanent file. However, the transferring school was not required to send the teacher copy of the checklist or the teacher file of observations and student work.

2.2 ADMINISTRATOR TRAINING

After the 2000 and 2001 training workshops, the trained district personnel became responsible for training new teachers. In addition, approximately ten thousand copies of the *South Carolina Readiness Assessment Kindergarten and First Grade Developmental Guidelines* (Rebus and SDE 2002) were distributed to teachers statewide in July and August 2002.

The guidelines prior to 2002–03 were rewritten to reflect changes in the WSS and the revised South Carolina English language arts curriculum standards. The new 2002–03 guidelines also combined kindergarten and the first grade in a single document for easy reference. Additionally, the standards associated with each indicator in mathematics and English language arts were listed beneath the indicator along with the rationale and examples. The same guidelines were used in 2003–04.

In September 2003, district personnel who participated in the pilot of the SCRAPI online system were trained in the use of the system in a workshop conducted by the SDE. In addition, the training was made available during a live broadcast over the Instructional Television network.

In training workshops conducted by the SDE in March 2004, district test coordinators received copies of the *South Carolina Readiness Assessment Developmental Checklists: District and School Test Administrator Manual 2003–2004* (HEM 2003). In April 2004, schools received their copies of the manual. Although the manual contains procedures for returning paper checklists, it also gives explanations for the student demographic coding that must be completed for both the paper and the online systems.

2.3 SOUTH CAROLINA READINESS ASSESSMENT PROFILE INTERACTIVE

The South Carolina Readiness Assessment Profile Interactive (SCRAPI) online system allows teachers to submit their completed checklists electronically. This section summarizes the findings of Porchea and Casteel (2003) regarding the system.

Initial implementation of the SCRAPI system during the 2002–03 school year was invaluable in generating feedback about the system. The feedback was used to enhance the user-friendliness and utility of the system for subsequent SCRA administration years.

History

Since the inception of the SCRA, teachers had been concerned about the amount of paperwork involved, and efforts were made to ease their concerns. The SCRA checklists are now shorter than the WSS checklists, from which they were derived. With advisory committee support, the SDE eased the burden in 2001–02 and 2002–03 by reducing the mandatory rating periods from three (fall, winter, and end of year) to two (winter and end of year) and providing preprinted labels to override the demographic page, which would have to be completed manually if they did not have the preprinted labels.

In 2001, the SDE contracted with the state's Horry County School District, which had already begun creating an online reporting system for its own use. The system was first tested by all 2001–02 kindergarten and first-grade teachers in the Horry County School District and in Dillon School District Two. Approximately 250 teachers completed the ratings on over 5,100 students.

Once the SCRAPI software was installed in the state computer system, student demographic data were loaded into the SCRAPI system. Districts submitted student, teacher, and parent information to the state at the beginning of the 2002–03 school year. The data were then loaded into the SCRAPI system. Hence, teachers logging on to the system for the first time found their classroom information already on the system.

District testing coordinators and technology coordinators were invited to volunteer their districts, as a whole or in part, to pilot SCRAPI in the 2002–03 school year. Twenty-two districts successfully participated in the pilot. Some of these districts used SCRAPI districtwide; others used SCRAPI in only some of their schools.

Online System Features

Since the teacher and student information was preloaded into SCRAPI, teachers logging on to the system for the first time enter minimal identifying information and an e-mail address. The system assigns a random password and sends it to the teacher via e-mail.

Once logged on to the system, the teachers have access to several menu options. Teachers complete online student checklists, which are designed to resemble paper checklists.

Teachers can also view a class summary showing the proportion of each student's checklist that has been completed as well as the proportion of the total class with completed checklists. Since the summary pages in SCRAPI access the state database, the summaries are always current. A summary of the total ratings is available (*rarely or never demonstrates, sometimes demonstrates,* or *consistently demonstrates*) for teachers to be able to assess their overall class levels at any given point in the year.

School administrators can view a teacher-ratings summary that shows where teachers are in the rating process (i.e., the percentage of checklists completed). School administrators can view individual student ratings but do not have privileges to add or modify student ratings.

District administrators can view a summary that shows the number of ratings completed at each of their schools. (A district administrator can also function as a school administrator in any of the schools in his or her district.) State administrators can view a summary of all the districts that shows the number of ratings completed in each district as well as the total number of students in the system. (A state administrator can also function as a district/school administrator for any district/school.)

Transferring Students

A common complaint made by school personnel had been that students transfer into their schools during the school year without all of the necessary documents from the previous school's files. Even when the sending school included all documents from the student's permanent records, often the SCRA checklist was not included because it may have been stored in the classroom rather than in the school office. Now, however, if a student is in the SCRAPI system, his or her ratings are already electronically stored and it is simply a matter of reassigning the student to the receiving school.

School administrators handle student transfers. SCRAPI has a feature that allows school administrators to search for students with a similar name or other identifying information. Once the student record is found and selected, if the record is marked *inactive* (meaning the student is not currently assigned to a school), the school administrator selects the student record and designates it to the new teacher. If the student record is marked *active*, the SCRAPI system automatically sends an e-mail to the school administrator at the sending school requesting that the student record be released. Once the sending school releases it, another e-mail is generated to notify the receiving school that the student record can now be transferred.

2.4 SCORING

At the end of the 2002–03 school year, the data collected through SCRAPI were sent to the HEM for processing. HEM consolidated the SCRAPI data with the SCRA data collected from paper checklists. Scoring of the 2002–03 checklists for kindergarten and grade one proceeded in the same manner as in previous years. Under contract to the SDE, the HEM was responsible for receiving, scanning, scoring, and reporting SCRA results. The HEM and the SDE determined the scoring algorithms and report formats for the readiness assessment. Consultants from Rebus, the HEM, and the SDE determined the format and content of score reports for students, districts, and the state.

Using the checklists, teachers made four assessments of their students' performance with regard to the curriculum standards: performance to date (fall), performance to date (winter), performance to date (spring), and year-end performance. Only year-end performance ratings were used in the reporting process. For paper checklists, number-two lead pencils were to be used to mark all ratings on the official copy. Schools then sent the completed scannable checklists to the district testing coordinator, who in turn sent them to the HEM San Antonio scoring center for scoring at the end of the school year.

The HEM's technical proposal to the SDE on the readiness assessment for first and second graders explains the quality control and scoring systems:

> Harcourt's WorkFlow System is designed to provide several checks/edits of scannable documents for accuracy. The WorkFlow System allows for a second edit to check for: 1) the completeness of information provided on scannable materials; 2) scannable documents returned in poor condition; and 3) missing students', schools' or districts' documents. Several workstations are involved in this process.
>
> The workstation executes the resolve/edit program and produces the edit list. The edit list is a sequential listing of error suspect records, including SSIDs. The suspect field(s) within a record are printed with a flag. This flag means that the data field did not pass our edit requirements. . . . In each case, an error suspect record must be reviewed by an editor who references the source answer document to ensure that the integrity of the data provided is maintained. Sequence numbers of the answer documents aid the editor in this review. . . .
>
> Score reports for the South Carolina Readiness Assessment will be printed only after an internal quality assurance and an external Department quality assurance check have verified accuracy. To further ensure accurate reporting for all districts, Harcourt will conduct a pre-print and post-print quality assurance check of every school, district, and building report. (HEM 1999, 45–46)

2.5 SCORE REPORTING

Score reporting for the 2003–04 kindergarten and grade-one data was the same as that for previous years. After the data were collected and other data analyses had been conducted, the individual student home reports; the class rosters; and the school, district, and state summaries were produced and shipped to the school districts. The district test coordinator was responsible

for ensuring that schools received the student home reports and other reports at the beginning of the 2004–05 school year.

Student Home Report

The 2003–04 student home report was a single 8½ x 11 inch page that used the performance levels *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates* to designate the student's strengths and needs in each of the functional components. Explanations of the performance levels and performance indicators were also provided in the student home report. The component rating was reported using an algorithm based on the scoring rules described in table 2.1 on the next page.

Class Roster

The class roster was an 8½ x 11 inch multipage document provided for every teacher by grade level. Students' names were listed in alphabetical order, and each student's performance was reported for each functional component.

School, District, and State Summaries

The school report was titled "Class List with School Summary"; the district report, "School List with District Summary," and the state report, "District List with State Summary." These 8½ x 11 inch multipage reports provided summary information for three subject domains called "Personal and Social Development," "English Language Arts," and "Mathematics." Performance-level information was reported by the number and the percentage of students at the *rarely or never demonstrates*, *sometimes demonstrates*, and *consistently demonstrates* levels for each of the functional components within the three domains. The last page of the report provided a summary of demographic data used on the checklists.

TABLE 2.1

Frequency of Performance Ratings Needed to Achieve Various Component Ratings

Number of Component Ratings	Frequency of Performance Ratings			Component Rating	
	Consistently Demonstrates	Sometimes Demonstrates	Rarely or Never Demonstrates	Title	Score
1	1	0	0	consistently demonstrates	3
1	0	1	0	sometimes demonstrates	2
1	0	0	1	rarely or never demonstrates	1
2	2	0	0	consistently demonstrates	3
2	1	1	0	sometimes demonstrates	2
2	1	0	1	sometimes demonstrates	2
2	0	2	0	sometimes demonstrates	2
2	0	1	1	sometimes demonstrates	2
2	0	0	2	rarely or never demonstrates	1
3	3	0	0	consistently demonstrates	3
3	2	1	0	consistently demonstrates	3
3	2	0	1	sometimes demonstrates	2
3	1	2	0	sometimes demonstrates	2
3	1	1	1	sometimes demonstrates	2
3	1	0	2	sometimes demonstrates	2
3	0	3	0	sometimes demonstrates	2
3	0	2	1	sometimes demonstrates	2
3	0	1	2	rarely or never demonstrates	1
3	0	0	3	rarely or never demonstrates	1
4	4	0	0	consistently demonstrates	3
4	3	1	0	consistently demonstrates	3
4	3	0	1	sometimes demonstrates	2
4	2	2	0	sometimes demonstrates	2
4	2	1	1	sometimes demonstrates	2
4	2	0	2	sometimes demonstrates	2
4	1	3	0	sometimes demonstrates	2
4	1	2	1	sometimes demonstrates	2
4	1	1	2	sometimes demonstrates	2
4	1	0	3	sometimes demonstrates	2
4	0	4	0	sometimes demonstrates	2
4	0	3	1	sometimes demonstrates	2
4	0	2	2	sometimes demonstrates	2
4	0	1	3	rarely or never demonstrates	1
4	0	0	4	rarely or never demonstrates	1
5	5	0	0	consistently demonstrates	3
5	4	1	0	consistently demonstrates	3
5	4	0	1	sometimes demonstrates	2
5	3	2	0	sometimes demonstrates	2
5	3	1	1	sometimes demonstrates	2
5	3	0	2	sometimes demonstrates	2
5	2	3	0	sometimes demonstrates	2
5	2	2	1	sometimes demonstrates	2
5	2	1	2	sometimes demonstrates	2
5	2	0	3	sometimes demonstrates	2
5	1	4	0	sometimes demonstrates	2
5	1	3	1	sometimes demonstrates	2
5	1	2	2	sometimes demonstrates	2
5	1	1	3	rarely or never demonstrates	1
5	1	0	4	rarely or never demonstrates	1
5	0	5	0	sometimes demonstrates	2
5	0	4	1	sometimes demonstrates	2
5	0	3	2	sometimes demonstrates	2
5	0	2	3	rarely or never demonstrates	1
5	0	1	4	rarely or never demonstrates	1
5	0	0	5	rarely or never demonstrates	1

Chapter 3

TECHNICAL CHARACTERISTICS OF THE INDICATORS AND RATINGS

As noted previously, the SCRA development checklists for 2003–04 were comprised of thirty-seven performance indicators for kindergarten and thirty-nine performance indicators for the first grade. Values for each indicator were assigned four times during the year: performance to date (fall), performance to date (winter), performance to date (spring), and year-end performance. Only the winter and year-end ratings were required. Values for the fourteen functional component ratings were derived from the year-end performance ratings. This section documents the technical characteristics of the indicators and ratings. For ease of terminology, these quantities are referred to globally as "items" in the remaining part of this chapter.

For each item, the SDE computed statistics such as the number of students, item mean, and item/total score correlation. Correlations were computed between the item and the total score (which includes that item). Appendix A provides these item statistics in detail for the year-end performance indicators and the year-end performance ratings in kindergarten and the first grade.

The 2003–04 statistics in tables 3.1 and 3.2 of this chapter include data from 2,160 students (of whom 1,011 were kindergarteners and 1,149 were first graders) who received alternate scoring as reported by the schools. Students who receive alternate scoring on the SCRA are rated only on the indicators that align with their individualized education plan (IEP) goals and objectives. The numbers of students who received alternate scoring were overreported by the schools. For future SCRA administrations, the SDE will caution the schools on this matter. Data on students who received alternate scoring are therefore excluded from tables 3.3 and 3.4 in this chapter.

3.1 YEAR-END PERFORMANCE INDICATORS

Table 3.1 gives a summary of the year-end performance indicators in statewide kindergarten and the first grade as well as for each gender (male and female) and ethnicity (African American and white) group. The effective sample size for each statistic varied since not all students had complete data.

TABLE 3.1

Summary of Major Statistics for Year-End Performance Indicators

	Student Group				
	All Students	Males	Females	African Americans	Whites
Kindergarten					
Number of students	46,493	24,084	22,359	18,146	25,189
Number of items	37	37	37	37	37
Average of item means	2.714	2.657	2.775	2.627	2.785
Average of item standard deviations	0.512	0.549	0.458	0.563	0.453
Average of correlations between items and total score	0.741	0.743	0.728	0.741	0.727
Alpha coefficient	0.978	0.978	0.977	0.978	0.976
Grade 1					
Number of students	47,986	24,849	23,065	18,687	26,193
Number of items	39	39	39	39	39
Average of item means	2.689	2.632	2.751	2.580	2.769
Average of item standard deviations	0.531	0.566	0.480	0.592	0.465
Average of correlations between items and total score	0.756	0.759	0.744	0.762	0.732
Alpha coefficient	0.981	0.981	0.980	0.982	0.977

Source: State Department of Education

3.2 YEAR-END PERFORMANCE RATINGS

Table 3.2 provides a summary of the year-end performance ratings in statewide kindergarten and the first grade as well as for each gender (male and female) and ethnicity (African American and white) group. The effective sample size for each statistic varied since not all students had complete data. Tables 3.3 and 3.4 provide the state summary since 2000–01 and illustrate the changes in performance ratings terminology since 2000–01.

TABLE 3.2

Summary of Major Statistics for Year-End Performance Ratings

	Student Group				
	All Students	Males	Females	African Americans	Whites
Kindergarten					
Number of students	47,468	24,714	22,701	18,627	25,631
Number of items	14	14	14	14	14
Average of item means	2.685	2.627	2.749	2.592	2.762
Average of item standard deviations	0.517	0.550	0.468	0.562	0.462
Average of correlations between items and total score	0.765	0.768	0.751	0.764	0.753
Alpha coefficient	0.950	0.951	0.945	0.950	0.946
Grade 1					
Number of students	49,112	25,588	23,452	19,216	26,748
Number of items	14	14	14	14	14
Average of item means	2.662	2.605	2.725	2.546	2.747
Average of item standard deviations	0.533	0.562	0.488	0.587	0.471
Average of correlations between items and total score	0.778	0.781	0.768	0.780	0.758
Alpha coefficient	0.953	0.954	0.951	0.954	0.946

Source: State Department of Education

TABLE 3.3

State Summary of Year-End Performance Ratings of the Kindergarten Assessment for First-Grade Readiness

Ratings	2002–03			2003–04		
	Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates	Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates
English Language Arts						
Communication	2%	32%	66%	2%	31%	67%
Reading	3%	27%	70%	3%	25%	72%
Writing	3%	21%	76%	3%	19%	78%
Mathematics						
Mathematical Processes	6%	35%	59%	6%	34%	60%
Number and Operations	2%	24%	74%	2%	22%	76%
Patterns, Relationships, and Functions	1%	21%	78%	1%	19%	79%
Geometry and Spatial Relations	1%	24%	75%	1%	23%	76%
Measurement	4%	29%	67%	3%	27%	70%
Data Collection and Probability	5%	31%	64%	5%	28%	67%
Personal and Social Development						
Self-Concept	2%	36%	62%	2%	36%	62%
Self-Control	2%	20%	78%	2%	20%	78%
Approaches to Learning	5%	29%	67%	4%	28%	68%
Interaction with Others	1%	21%	79%	1%	19%	80%
Social Problem-Solving	3%	24%	73%	3%	22%	75%

Note: For each school year, the sum of the percentages may equal 99 percent or 101 percent due to rounding to the nearest whole percentage.

TABLE 3.4

State Summary of Year-End Performance Ratings of the First-Grade Assessment for Second-Grade Readiness

Ratings	2002–03			2003–04		
	Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates	Rarely or Never Demonstrates	Sometimes Demonstrates	Consistently Demonstrates
English Language Arts						
Communication	3%	35%	62%	3%	33%	64%
Reading	4%	29%	67%	3%	27%	69%
Writing	6%	29%	65%	6%	27%	67%
Mathematics						
Mathematical Processes	3%	33%	64%	3%	31%	67%
Number and Operations	3%	21%	75%	3%	20%	77%
Patterns, Relationships, and Functions	1%	20%	78%	1%	18%	80%
Geometry and Spatial Relations	2%	31%	67%	2%	28%	71%
Measurement	3%	27%	70%	3%	25%	72%
Data Collection and Probability	4%	34%	63%	3%	31%	66%
Personal and Social Development						
Self-Concept	3%	38%	59%	3%	37%	60%
Self-Control	2%	21%	77%	2%	21%	77%
Approaches to Learning	5%	27%	68%	5%	26%	69%
Interaction with Others	1%	23%	76%	1%	22%	77%
Social Problem-Solving	5%	30%	64%	5%	29%	66%

Note: For each school year, the sum of the percentages may equal 99 percent or 101 percent due to rounding to the nearest whole percentage.

Chapter 4

VALIDITY

This chapter reports on measuring fairness using differential item functioning (DIF) for categories of gender and ethnicity.

4.1 OVERALL VIEW OF DIF ANALYSIS

General Description

One threat to the validity of an assessment is test bias, or the unfair advantage of one group over another on the test. DIF occurs, in the words of Clauser and Mazor, when "examinees from different groups have differing probabilities or likelihoods of success on an item, after they have been matched" on a characteristic of interest such as ability or achievement on the test as a whole. "DIF is a necessary but not sufficient condition for item bias," they explain (Clauser and Mazor 1998, 31). DIF focuses on *item* validity as opposed to *test* validity because the procedures assume that the test as a whole is a "good" measure.

For documentation purposes, SCRA items (performance indicators and performance ratings) were subjected to a formal DIF analysis based on the Mantel-Haenszel (MH) procedure, which has a long tradition in DIF analysis and is considered effective and efficient (see Clauser and Mazor 1998, Hills 1989). The MH procedure uses both a statistical significance test and an analysis of the effect size.

In the MH procedure, total scores (on the entire scale) are typically used to group students into strata. Students in each stratum were considered as equivalent in terms of the construct under assessment. Then, for each item, the students in the "focal" and "reference" groups were compared on the basis of their mean score on the item. The term *focal* refers to the group of interest for DIF—in this case, female or African American. The comparison or reference group was male or white, depending upon whether the DIF analysis was for gender or ethnicity.

In traditional National Assessment of Educational Progress (NAEP) DIF analysis for multiple-choice items, items are classified as "A," "B," or "C" in terms of DIF. "A" items are considered to be free of DIF. "B" items may be used, but if there is a choice among otherwise equivalent items, it is considered desirable to select for inclusion in a test those items with the smallest absolute value of MH DIF. "C" items are to be selected only if they are essential to meet test specifications (Zwick and Ercikan 1989, 58–59). For other items with more than two ordered response categories (such as SCRA items), the corresponding NAEP DIF classifications are "AA," "BB," and "CC."

Standardized Mean Difference Procedure

For the 2002–03 year-end performance SCRA data, the MH procedure for polytomous items was used. This procedure is based on the Cochran-Mantel-Haenszel (CMH) procedure for ordered categorical variables and an analysis of the standardized mean difference (SMD). (See Dorans, Schmitt, and Bleistein 1992; Zwick and Thayer 1996; Zwick, Thayer, and Mazzeo 1997.) Values of the SMD serve as measures of effect size for constructed-response items.

The MH procedure was implemented using SAS (Stokes, Davis, and Koch 2001). Students were grouped into ten strata on the basis of their complete total raw scores. The FREQ procedure provided the MH statistic with a *p*-value. The standard deviation (SD) of the combined group, the unweighted mean (MF) of the focal group, and the weighted mean (MR) of the reference group were computed using SAS. The MR was obtained by taking the mean of the reference group in each stratum and weighting it according to the number of focal students in the same stratum. The MR subtracted from the MF is the SMD. An estimate of the effect size (ES) of the mean was then obtained by dividing the SMD by the SD. Following the recommendation of the NAEP (see Allen, Carlson, and Zalanak 1999), the combined group SD was used as a scale measure for the item. With several DIF analyses to be carried out for the same set of data, it was necessary to keep the scale constant so that all ESs *for the same item* would be expressed on a common scale and therefore compatible with each other across different DIF analyses.

NAEP DIF Classification

The following ETS/NAEP rules were used to classify items as "AA," "BB," and "CC."

Rule 1: If the *p*-value > 0.05, classify the item as AA.

Otherwise, use the following rules based on the absolute value of the ES:

Rule 2: If abs(ES) is less than or equal to 0.17, classify the item as AA.
Rule 3: If abs(ES) is greater than 0.17 and less than or equal to 0.25, classify the item as BB.
Rule 4: If abs(ES) is greater than 0.25, then classify the item as CC.

4.2 RESULTS OF DIF ANALYSES

Table 4.1 provides a summary of DIF classifications for the year-end performance indicators for both kindergarten and grade-one students. Similar data for the year-end performance ratings are listed in table 4.2. For the year-end performance indicators, the total 76 indicators were classified as "AA" for 74 cases and "BB" for 2 cases in gender DIF. All 76 cases were classified as "AA" in ethnicity DIF. For the year-end performance ratings, the total 28 ratings were classified as "AA" for 27 cases and "BB" for 1 case in gender DIF. All 28 ratings were classified as "AA" in ethnicity DIF. The data clearly indicate that the checklists are reasonably free of gender and ethnicity DIF. Appendix B provides complete DIF data for year-end performance indicators and year-end performance ratings. The DIF analyses also include data from students who received alternate scoring as reported by the schools.

TABLE 4.1

Summary of DIF Classification for Year-End Performance Indicators

Reference Group	Focal Group	Total Number of Items	NAEP DIF Classification		
			AA	BB	CC
Kindergarten					
Males	Females	37	36	1	---
Whites	African Americans	37	37	---	---
First Grade					
Males	Females	39	38	1	---
Whites	African Americans	39	39	---	---
Both Grades					
Males	Females	76	74	2	---
Whites	African Americans	76	76	---	---

TABLE 4.2

Summary of DIF Classification for Year-End Performance Ratings

Reference Group	Focal Group	Total Number of Items	NAEP DIF Classification		
			AA	BB	CC
Kindergarten					
Males	Females	14	14	---	---
Whites	African Americans	14	14	---	---
First Grade					
Males	Females	14	13	1	---
Whites	African Americans	14	14	---	---
Both Grades					
Males	Females	28	27	1	---
Whites	African Americans	28	28	---	---

Appendix A

ITEM STATISTICS FOR YEAR-END PERFORMANCE INDICATORS AND YEAR-END PERFORMANCE RATINGS

Provided here are the following tables with detailed DIF analysis results for gender and ethnicity in kindergarten and the first grade:

Table A1: Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade
Table A2: Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade

The following terms are used in these two tables:

Indicator	=	item sequence as it appears in the checklist
Functional Component	=	functional component name
Item Mean	=	arithmetic mean of the item (Responses were assigned the following weights: 1 for *rarely or never demonstrates*, 2 for *sometimes demonstrates*, and 3 for *consistently demonstrates*. The mean includes all students with a rating on the item regardless of whether they have complete ratings on all other items.)
Item-Scale Correlation	=	correlation between the item and the total score on the scale (Only records with responses to all the checklist items were included in the correlation. The term *scale* denotes the total score of all items that form the checklist.)
N	=	number of students with responses to the item

TABLE A.1

Item Statistics: Year-End Performance Indicators for Kindergarten and First Grade

Kindergarten				First Grade			
Indicator	**Item Mean**	**N**	**Item-Scale Correlation**	**Indicator**	**Item Mean**	**N**	**Item-Scale Correlation**
1	2.665	47,394	0.693	1	2.632	49,024	0.727
2	2.618	47,366	0.761	2	2.578	49,012	0.787
3	2.669	47,384	0.553	3	2.658	49,026	0.592
4	2.806	47,344	0.604	4	2.780	48,977	0.641
5	2.753	47,376	0.644	5	2.732	49,030	0.676
6	2.731	47,367	0.757	6	2.708	49,001	0.770
7	2.578	47,337	0.771	7	2.592	48,994	0.764
8	2.584	47,361	0.799	8	2.568	49,011	0.818
9	2.800	47,368	0.586	9	2.750	49,012	0.613
10	2.850	47,386	0.581	10	2.787	48,987	0.621
11	2.775	47,356	0.687	11	2.758	48,996	0.716
12	2.773	47,364	0.592	12	2.752	49,004	0.612
13	2.718	47,376	0.667	13	2.597	49,007	0.741
14	2.685	47,325	0.782	14	2.666	48,995	0.779
15	2.643	47,308	0.794	15	2.589	48,980	0.810
16	2.682	47,304	0.753	16	2.653	48,947	0.766
17	2.644	47,297	0.793	17	2.596	48,951	0.804
18	2.795	47,302	0.775	18	2.789	48,968	0.735
19	2.768	47,284	0.811	19	2.796	48,929	0.776
20	2.684	47,280	0.795	20	2.683	48,927	0.785
21	2.667	47,283	0.784	21	2.615	48,877	0.794
22	2.718	47,276	0.814	22	2.666	48,928	0.807
23	2.745	47,270	0.800	23	2.585	48,925	0.829
24	2.728	47,259	0.794	24	2.612	48,922	0.792
25	2.716	47,309	0.825	25	2.580	48,952	0.820
26	2.555	47,302	0.813	26	2.703	48,938	0.784
27	2.567	47,284	0.821	27	2.628	48,934	0.808
28	2.763	47,293	0.791	28	2.782	48,928	0.779
29	2.734	47,263	0.812	29	2.736	48,898	0.803
30	2.798	47,281	0.773	30	2.572	48,932	0.801
31	2.805	47,309	0.765	31	2.801	48,882	0.782
32	2.804	47,241	0.765	32	2.813	48,921	0.767
33	2.760	47,275	0.781	33	2.787	48,892	0.780
34	2.675	47,282	0.781	34	2.669	48,948	0.794
35	2.590	47,285	0.765	35	2.655	48,927	0.791
36	2.687	47,277	0.772	36	2.685	48,910	0.786
37	2.608	47,313	0.758	37	2.696	48,947	0.776
				38	2.679	48,942	0.806
				39	2.616	48,959	0.808

TABLE A.2

Item Statistics: Year-End Performance Ratings for Kindergarten and First Grade

	Kindergarten			First Grade		
Functional Component	**Item Mean**	**N**	**Item-Scale Correlation**	**Item Mean**	**N**	**Item-Scale Correlation**
Self-Concept	2.593	47,326	0.715	2.563	48,960	0.743
Self-Control	2.762	47,246	0.628	2.744	48,883	0.664
Approaches to Learning	2.628	47,229	0.811	2.634	48,876	0.826
Interaction with Others	2.793	47,185	0.664	2.749	48,807	0.688
Social Problem-Solving	2.718	47,376	0.667	2.597	49,007	0.741
Communication	2.638	47,133	0.814	2.600	48,820	0.816
Reading	2.678	46,993	0.834	2.652	48,627	0.815
Writing	2.741	47,142	0.831	2.598	48,817	0.836
Mathematical Processes	2.539	47,229	0.807	2.629	48,861	0.784
Number and Operations	2.728	47,183	0.803	2.736	48,754	0.814
Patterns, Relationships, and Functions	2.772	47,220	0.783	2.785	48,798	0.785
Geometry and Spatial Relations	2.741	47,166	0.784	2.684	48,818	0.776
Measurement	2.656	47,124	0.812	2.685	48,770	0.804
Data Collection and Probability	2.608	47,313	0.758	2.613	48,908	0.795

Appendix B

DIFFERENTIAL ITEM FUNCTIONING FOR YEAR-END PERFORMANCE INDICATORS AND YEAR-END PERFORMANCE RATINGS

Provided here are the following tables with detailed DIF analysis results for gender and ethnicity in kindergarten and the first grade:

Table B.1 DIF Data: Year-End Performance Indicators for Kindergarten and Gender
Table B.2 DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity
Table B.3 DIF Data: Year-End Performance Indicators for First Grade and Gender
Table B.4 DIF Data: Year-End Performance Indicators for First Grade and Ethnicity
Table B.5 DIF Data: Year-End Performance Ratings for Kindergarten and Gender
Table B.6 DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity
Table B.7 DIF Data: Year-End Performance Ratings for First Grade and Gender
Table B.8 DIF Data: Year-End Performance Ratings for First Grade and Ethnicity

The following terms are used in the tables:

Indicator	=	item sequence as it appears in the checklist
Functional Component	=	functional component name
p-value	=	MH statistic's probability for statistical significance (With a large number of students in the DIF groups, the *p*-value can be very small and may be indicated by 0.000, which means the probability is less than 0.0005.)
Focal Mean	=	unweighted mean of the focal group
Reference Mean	=	weighted mean of the reference group where weights are the proportion of focal students in each stratum over all the students in the focal group
SD	=	standard deviation of the item/rating over all focal and reference students
Effect Size	=	the SMD (the focal mean minus the reference mean) divided by the standard deviation
DIF	=	classification based on NAEP rules for differential item functioning [plus sign (+) = in favor of the focal group; minus sign (-) = against the focal group]

TABLE B.1

DIF Data: Year-End Performance Indicators for Kindergarten and Gender

Indicator	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.023	2.727	2.736	0.527	-0.016	AA
2	0.000	2.711	2.686	0.573	0.044	AA
3	0.000	2.784	2.681	0.529	0.194	BB
4	0.000	2.889	2.821	0.430	0.159	AA
5	0.000	2.838	2.784	0.481	0.113	AA
6	0.449	2.797	2.795	0.497	0.005	AA
7	0.000	2.687	2.646	0.605	0.068	AA
8	0.000	2.684	2.665	0.598	0.031	AA
9	0.000	2.856	2.837	0.437	0.042	AA
10	0.004	2.891	2.884	0.380	0.018	AA
11	0.000	2.836	2.824	0.465	0.025	AA
12	0.000	2.844	2.805	0.463	0.086	AA
13	0.000	2.798	2.764	0.505	0.069	AA
14	0.005	2.762	2.753	0.520	0.017	AA
15	0.000	2.736	2.711	0.559	0.046	AA
16	0.000	2.761	2.748	0.545	0.024	AA
17	0.000	2.716	2.731	0.580	-0.026	AA
18	0.000	2.861	2.841	0.443	0.045	AA
19	0.841	2.829	2.830	0.480	-0.002	AA
20	0.674	2.759	2.758	0.564	0.002	AA
21	0.442	2.744	2.742	0.571	0.004	AA
22	0.000	2.774	2.797	0.519	-0.044	AA
23	0.000	2.818	2.804	0.496	0.029	AA
24	0.000	2.807	2.787	0.516	0.038	AA
25	0.008	2.792	2.785	0.532	0.013	AA
26	0.000	2.612	2.675	0.618	-0.103	AA
27	0.000	2.624	2.686	0.616	-0.099	AA
28	0.000	2.804	2.838	0.486	-0.069	AA
29	0.000	2.780	2.817	0.511	-0.072	AA
30	0.000	2.840	2.861	0.448	-0.047	AA
31	0.000	2.848	2.865	0.446	-0.038	AA
32	0.000	2.841	2.866	0.438	-0.056	AA
33	0.000	2.805	2.832	0.476	-0.056	AA
34	0.000	2.724	2.766	0.537	-0.079	AA
35	0.000	2.641	2.697	0.579	-0.097	AA
36	0.000	2.739	2.771	0.540	-0.059	AA
37	0.000	2.671	2.704	0.585	-0.057	AA

TABLE B.2

DIF Data: Year-End Performance Indicators for Kindergarten and Ethnicity

Indicator	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.000	2.601	2.554	0.526	0.090	AA
2	0.000	2.531	2.488	0.573	0.075	AA
3	0.000	2.563	2.613	0.533	-0.094	AA
4	0.000	2.734	2.753	0.433	-0.045	AA
5	0.047	2.679	2.671	0.484	0.016	AA
6	0.001	2.650	2.637	0.498	0.024	AA
7	0.000	2.469	2.441	0.607	0.047	AA
8	0.050	2.467	2.459	0.600	0.013	AA
9	0.000	2.731	2.749	0.439	-0.043	AA
10	0.194	2.801	2.807	0.379	-0.016	AA
11	0.000	2.717	2.693	0.464	0.053	AA
12	0.000	2.687	2.730	0.465	-0.093	AA
13	0.000	2.620	2.652	0.505	-0.063	AA
14	0.000	2.598	2.583	0.520	0.028	AA
15	0.000	2.549	2.529	0.559	0.036	AA
16	0.030	2.602	2.602	0.536	0.000	AA
17	0.000	2.537	2.570	0.570	-0.058	AA
18	0.000	2.738	2.714	0.439	0.054	AA
19	0.050	2.697	2.688	0.475	0.019	AA
20	0.000	2.599	2.585	0.558	0.026	AA
21	0.000	2.583	2.562	0.566	0.036	AA
22	0.436	2.634	2.638	0.511	-0.009	AA
23	0.003	2.671	2.658	0.491	0.027	AA
24	0.000	2.655	2.631	0.512	0.046	AA
25	0.100	2.630	2.623	0.528	0.013	AA
26	0.000	2.424	2.451	0.614	-0.044	AA
27	0.000	2.444	2.463	0.610	-0.032	AA
28	0.010	2.690	2.679	0.482	0.023	AA
29	0.127	2.650	2.651	0.507	-0.003	AA
30	0.464	2.730	2.734	0.442	-0.008	AA
31	0.006	2.734	2.745	0.441	-0.025	AA
32	0.579	2.744	2.741	0.430	0.008	AA
33	0.000	2.683	2.703	0.466	-0.043	AA
34	0.000	2.572	2.600	0.532	-0.052	AA
35	0.000	2.475	2.503	0.575	-0.049	AA
36	0.033	2.605	2.595	0.534	0.018	AA
37	0.000	2.498	2.522	0.580	-0.040	AA

Table B.3

DIF Data: Year-End Performance Indicators for First Grade and Gender

Indicator	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.280	2.698	2.703	0.548	-0.008	AA
2	0.000	2.680	2.644	0.604	0.061	AA
3	0.000	2.779	2.670	0.539	0.201	BB
4	0.000	2.871	2.798	0.459	0.159	AA
5	0.000	2.825	2.761	0.497	0.129	AA
6	0.000	2.786	2.764	0.519	0.042	AA
7	0.000	2.720	2.636	0.607	0.138	AA
8	0.000	2.670	2.644	0.615	0.043	AA
9	0.000	2.811	2.794	0.480	0.036	AA
10	0.000	2.851	2.820	0.444	0.070	AA
11	0.000	2.820	2.808	0.480	0.025	AA
12	0.000	2.830	2.782	0.486	0.099	AA
13	0.000	2.700	2.654	0.585	0.078	AA
14	0.000	2.749	2.723	0.533	0.049	AA
15	0.000	2.695	2.654	0.599	0.068	AA
16	0.000	2.735	2.713	0.561	0.038	AA
17	0.060	2.679	2.677	0.604	0.004	AA
18	0.000	2.863	2.823	0.451	0.088	AA
19	0.000	2.854	2.844	0.457	0.023	AA
20	0.015	2.755	2.749	0.551	0.010	AA
21	0.000	2.697	2.691	0.600	0.011	AA
22	0.017	2.735	2.741	0.563	-0.011	AA
23	0.000	2.697	2.650	0.614	0.077	AA
24	0.000	2.707	2.675	0.594	0.054	AA
25	0.000	2.692	2.644	0.622	0.076	AA
26	0.000	2.726	2.798	0.526	-0.136	AA
27	0.000	2.653	2.742	0.581	-0.153	AA
28	0.000	2.804	2.857	0.463	-0.113	AA
29	0.000	2.761	2.823	0.507	-0.123	AA
30	0.000	2.611	2.689	0.606	-0.129	AA
31	0.000	2.831	2.865	0.441	-0.077	AA
32	0.000	2.843	2.872	0.431	-0.067	AA
33	0.000	2.816	2.856	0.451	-0.088	AA
34	0.000	2.706	2.763	0.548	-0.104	AA
35	0.000	2.692	2.753	0.548	-0.111	AA
36	0.000	2.720	2.777	0.532	-0.108	AA
37	0.000	2.728	2.786	0.522	-0.110	AA
38	0.000	2.722	2.769	0.544	-0.086	AA
39	0.000	2.657	2.723	0.584	-0.112	AA

Table B.4

DIF Data: Year-End Performance Indicators for First Grade and Ethnicity

Indicator	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
1	0.000	2.546	2.502	0.548	0.080	AA
2	0.000	2.469	2.424	0.605	0.075	AA
3	0.000	2.526	2.602	0.544	-0.140	AA
4	0.000	2.683	2.726	0.464	-0.092	AA
5	0.177	2.642	2.641	0.501	0.001	AA
6	0.000	2.611	2.599	0.521	0.023	AA
7	0.000	2.493	2.434	0.610	0.097	AA
8	0.836	2.430	2.432	0.616	-0.002	AA
9	0.000	2.656	2.686	0.484	-0.062	AA
10	0.001	2.710	2.729	0.445	-0.042	AA
11	0.007	2.679	2.670	0.480	0.019	AA
12	0.000	2.639	2.713	0.489	-0.152	AA
13	0.000	2.454	2.501	0.587	-0.080	AA
14	0.000	2.567	2.549	0.534	0.035	AA
15	0.000	2.472	2.449	0.600	0.039	AA
16	0.000	2.541	2.562	0.556	-0.039	AA
17	0.000	2.458	2.501	0.599	-0.071	AA
18	0.000	2.727	2.691	0.451	0.079	AA
19	0.000	2.726	2.711	0.453	0.033	AA
20	0.000	2.597	2.567	0.546	0.054	AA
21	0.000	2.515	2.481	0.595	0.057	AA
22	0.054	2.561	2.556	0.558	0.009	AA
23	0.000	2.474	2.437	0.610	0.061	AA
24	0.000	2.525	2.465	0.591	0.101	AA
25	0.000	2.470	2.430	0.618	0.064	AA
26	0.000	2.588	2.609	0.526	-0.039	AA
27	0.000	2.491	2.519	0.580	-0.050	AA
28	0.000	2.689	2.706	0.463	-0.038	AA
29	0.000	2.625	2.651	0.506	-0.052	AA
30	0.000	2.428	2.459	0.605	-0.052	AA
31	0.000	2.711	2.737	0.440	-0.059	AA
32	0.000	2.728	2.747	0.431	-0.044	AA
33	0.000	2.696	2.719	0.449	-0.051	AA
34	0.000	2.541	2.578	0.546	-0.068	AA
35	0.000	2.529	2.566	0.546	-0.068	AA
36	0.000	2.562	2.602	0.531	-0.076	AA
37	0.000	2.589	2.605	0.520	-0.030	AA
38	0.000	2.556	2.586	0.543	-0.056	AA
39	0.000	2.478	2.514	0.581	-0.062	AA

Table B.5

DIF Data: Year-End Performance Ratings for Kindergarten and Gender

Functional Component	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.002	2.673	2.662	0.536	0.020	AA
Self-Control	0.000	2.858	2.780	0.464	0.168	AA
Approaches to Learning	0.000	2.722	2.702	0.568	0.034	AA
Interaction with Others	0.000	2.852	2.834	0.420	0.042	AA
Social Problem-Solving	0.000	2.798	2.764	0.505	0.069	AA
Communication	0.005	2.721	2.713	0.527	0.015	AA
Reading	0.556	2.752	2.753	0.530	-0.001	AA
Writing	0.000	2.816	2.802	0.504	0.028	AA
Mathematical Processes	0.000	2.593	2.659	0.604	-0.109	AA
Number and Operations	0.000	2.771	2.810	0.494	-0.080	AA
Patterns, Relationships, and Functions	0.000	2.815	2.840	0.449	-0.054	AA
Geometry and Spatial Relations	0.000	2.785	2.815	0.464	-0.066	AA
Measurement	0.000	2.709	2.752	0.547	-0.078	AA
Data Collection and Probability	0.000	2.671	2.704	0.585	-0.057	AA

Table B.6

DIF Data: Year-End Performance Ratings for Kindergarten and Ethnicity

Functional Component	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.521	2.466	0.536	0.104	AA
Self-Control	0.000	2.677	2.700	0.467	-0.048	AA
Approaches to Learning	0.000	2.520	2.504	0.569	0.029	AA
Interaction with Others	0.003	2.722	2.735	0.421	-0.030	AA
Social Problem-Solving	0.000	2.620	2.652	0.505	-0.063	AA
Communication	0.223	2.544	2.540	0.524	0.007	AA
Reading	0.000	2.592	2.576	0.525	0.032	AA
Writing	0.001	2.661	2.650	0.499	0.023	AA
Mathematical Processes	0.000	2.411	2.436	0.600	-0.043	AA
Number and Operations	0.049	2.645	2.643	0.491	0.003	AA
Patterns, Relationships, and Functions	0.020	2.697	2.705	0.444	-0.017	AA
Geometry and Spatial Relations	0.000	2.664	2.675	0.456	-0.025	AA
Measurement	0.000	2.552	2.570	0.542	-0.032	AA
Data Collection and Probability	0.000	2.498	2.522	0.580	-0.040	AA

Table B.7

DIF Data: Year-End Performance Ratings for First Grade and Gender

Functional Component	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.649	2.629	0.551	0.037	AA
Self-Control	0.000	2.846	2.763	0.480	0.171	BB
Approaches to Learning	0.000	2.738	2.694	0.572	0.078	AA
Interaction with Others	0.000	2.819	2.790	0.457	0.064	AA
Social Problem-Solving	0.000	2.700	2.654	0.585	0.078	AA
Communication	0.000	2.691	2.665	0.548	0.047	AA
Reading	0.001	2.729	2.721	0.545	0.014	AA
Writing	0.000	2.707	2.661	0.608	0.075	AA
Mathematical Processes	0.000	2.650	2.739	0.543	-0.163	AA
Number and Operations	0.000	2.762	2.823	0.506	-0.121	AA
Patterns, Relationships, and Functions	0.000	2.815	2.852	0.439	-0.084	AA
Geometry and Spatial Relations	0.000	2.714	2.772	0.500	-0.116	AA
Measurement	0.000	2.720	2.777	0.529	-0.107	AA
Data Collection and Probability	0.000	2.653	2.716	0.555	-0.113	AA

Table B.8

DIF Data: Year-End Performance Ratings for First Grade and Ethnicity

Functional Component	*P*-Value	Focal Group Mean	Reference Group Mean	Standard Deviation	Effect Size	DIF Classification
Self-Concept	0.000	2.467	2.427	0.552	0.072	AA
Self-Control	0.000	2.639	2.679	0.484	-0.082	AA
Approaches to Learning	0.000	2.519	2.497	0.574	0.037	AA
Interaction with Others	0.000	2.654	2.681	0.459	-0.059	AA
Social Problem-Solving	0.000	2.454	2.501	0.587	-0.080	AA
Communication	0.001	2.483	2.487	0.546	-0.007	AA
Reading	0.000	2.557	2.531	0.541	0.046	AA
Writing	0.000	2.491	2.446	0.604	0.074	AA
Mathematical Processes	0.000	2.502	2.526	0.543	-0.044	AA
Number and Operations	0.000	2.624	2.649	0.505	-0.051	AA
Patterns, Relationships, and Functions	0.000	2.692	2.717	0.438	-0.057	AA
Geometry and Spatial Relations	0.000	2.565	2.599	0.500	-0.068	AA
Measurement	0.000	2.564	2.596	0.528	-0.062	AA
Data Collection and Probability	0.000	2.482	2.515	0.554	-0.060	AA

WORKS CITED

Allen, Nancy L., James E. Carlson, and Christine A. Zalanak. 1999. *The NAEP 1996 Technical Report.* Washington, DC: National Center for Education Statistics.

Clauser, Brian E., and Kathleen M. Mazor. 1998. "An NCME Instructional Module on Using Statistical Procedures to Identify Differentially Functioning Test Items." *Educational Measurement: Issues and Practice* 17, no. 1 (spring): 31–44.

Dorans, Neil J., Alicia P. Schmitt, and Carole A. Bleistein. 1992. "The Standardization Approach to Assessing Comprehensive Differential Item Functioning." *Journal of Educational Measurement* 29:309–19.

HEM. 1999. *A Proposal to South Carolina Department of Education for the Readiness Assessment for First and Second Graders*. San Antonio, TX: Harcourt Educational Measurement.

———. 2003. *South Carolina Readiness Assessment Developmental Checklists: District and School Test Administrator Manual 2003–2004*. San Antonio, TX: Harcourt Educational Measurement.

Hills, J. R. 1989. "Screening for Potentially Biased Items in Testing Programs." *Educational Measurements Issues and Practice* 8, no. 4:5–11.

Huynh, Huynh, Susan V. Prior, and Dorinda J. Gallant-Taylor. 2004. *Technical Documentation for the 2002 South Carolina Readiness Assessment of Kindergarten and Grade-One Students.* Columbia: South Carolina Department of Education.

Monrad, Diane M., and Garrett K. Mandeville. 1996. *Evaluation of the Work Sampling System Field Test.* Columbia: South Carolina Educational Policy Center.

Porchea, Sameano F., and James F. Casteel. 2003. "The South Carolina Readiness Assessment Profile Interactive (SCRAPI)." Paper presented at annual meeting of the National Council on Measurement in Education. Chicago, IL, April 22–24.

Rebus and SDE. 2002. South Carolina Readiness Assessment Kindergarten and First Grade Developmental Guidelines. New York: Pearson Early Learning.

Stokes, Maura E., Charles S. Davis, and Gary G. Koch. 2001. *Categorical Data Analysis Using the SAS System.* 2nd ed. New York: Wiley.

Zwick, Rebecca, and Kadriye Ercikan. 1989. "Analysis of Differential Item Functioning in the NAEP History Assessment." *Journal of Educational Measurement* 26:55–66.

Zwick, Rebecca, and Dorothy T. Thayer. 1996. "Evaluating the Magnitude of Differential Item Functioning in Polytomous Items." *Journal of Educational and Behavioral Statistics* 21:187–201.

Zwick, Rebecca, Dorothy T. Thayer, and John Mazzeo. 1997. “Descriptive and Inferential Procedures for Assessing DIF in Polytomous Items.” *Applied Measurement in Education* 10:321–44.